Deep Learning-Inspired Image Quality Enhancement

Ruxin Wang

Faculty of Engineering and Information Technology

University of Technology Sydney

A thesis submitted for the degree of

Doctor of Philosophy

2017

To my loving parents
Jianxia Liu and *Jigang Wang*

Certificate of Original Authorship

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Ruxin Wang

Acknowledgements

I would like to sincerely thank everyone who has helped me to finish my doctoral studies.

First of all, I would like to express my sincere appreciation and deep gratitude to my supervisor **Prof. Dacheng Tao**. He has given me trust and freedom to pursue my research interests, and provided constructive suggestions to help me out of difficulties. I can always benefit and learn a lot from various detailed discussions with him, and be excited and energised by his amazing insight, unlimited patience, generous support, and constant encouragement. I feel very lucky to have had him as my supervisor.

I also wish to express my sincere appreciation to **Prof. Xinge You** who was my advisor when I was in Huazhong University of Science and Technology for master study. I am so grateful for his guidance and support. I would have not come to the research field of image processing and computer vision, and would not have met Prof. Dacheng Tao and come to UTS without him.

I would also like to give special thanks to my excellent collaborators: Prof. Jun Yu, Prof. Richang Hong, Dr. Kun Zeng, and Dr. Xiaoyan Li for their brilliant work and timely support. I have also been fortunate to work and have discussions with many other brilliant researchers: Dr Chang Xu, Dr Nannan Wang, Dr Fei Gao, Assistant Professor Chaohui Wang, Prof. Chen Gong, Dr Yong Luo, Dr Tao Liu, Prof. Kaibing Zhang, Dr Lianyang Ma, A/Prof Chenping Hou, Dr Xiao Liu, Dr Weilong Hou, A/Prof Bo Du, Prof Shengzheng Wang, A/Prof Tao Lei, A/Prof Wankou Yang, A/Prof Shigang Liu, A/Prof Xianye Ben, A/Prof Xiong Wang, and Dr Qiong Wang. I wish to

express appreciation to all of them for their support and kind companion.

I am so grateful to my colleagues and friends I met in Sydney: Mingming Gong, Tongliang Liu, Shaoli huang, Qiang Li, Zhibin Hong, Meng Fang, Changxing Ding, Maoying Qiao, Zhe Xu, Long Lan, Wei Bian, Tianyi Zhou, Jun Li, Chunyang Liu, Bozhong Liu, Zhiguo Long, Guodong Long, Jing Jiang, Huan Fu, Baosheng Yu, Zhe Chen, Zijing Chen, Xiyu Yu, Liu Liu, Yali Du, Guoliang Kang, Hao Xiong, Jiayan Qiu, Jiankang Deng, Chaoyue Wang, Dayong Tian, Jiang Bian, Bo Han, Kede Ma, Tianrong Rao, Lingxiang Wu, Sujuan Hou, Xiaoqing Yin, Peicheng Zhou, Mingjin Zhang, Tao Zhang, Haifeng Liu, Peng Hao, Haishuai Wang, and Xun Yang. All my friends here have provided strong support during both happy and stressful time. I am also incredibly grateful to my old friends: Chengbin Yan, Digang Wang, Xinwei Liu, Zhe Wang, Hao Ding, Junge Shen, Xin Wang, Ting Yuan, Haoran Lv, and Yuannan Zhao, who have accompanied me from my bachelor's degree to doctoral studies. I owe my deepest thanks to all of them!

Finally, I would like to express deep-felt gratitude to my family: my parents, my grandparents, my uncles and aunties, and my cousins, for their endless love, trust, encouragement, and full support throughout my studies and life.

I dedicate this thesis to them.

Abstract

Enhancing image quality is a classical image processing problem that has received plenty of attention over the past several decades. A high-quality image is always expected in various vision tasks, and degradations such as noise, low-resolution, and blur are required to be removed. While the conventional techniques for this task have achieved great progress, the recent top performer, deep models, can substantially and significantly boost performance compared with conventional ones. The advantages of deep learning which enables it to achieve such success are its high representational capacity and the strong nonlinearity of the models. In this thesis, we explore the development of advanced deep models for image quality enhancement by researching several fundamental issues with different motivations.

In particular, we are first motivated by a pivotal property of the human perceptual system that similar visual cues can stimulate the same neuron to induce similar neurological signals. However, image degradations can result in the fact that similar local structures in images exhibiting dissimilar observations. While the conventional neural networks do not consider this important property, we develop the (stacked) non-local auto-encoder which exploits self-similar information in natural images for enhancing the stability of signal propagation in the network. It is expected that similar structures should induce similar network propagation. This is achieved by constraining the difference between the hidden representations of non-local similar image blocks during training. By applying the proposed model to image restoration, we then develop a "collaborative stabilisation" step to further rectify forward propagation.

When applying deep models to image quality enhancement tasks, we are concerned about which factor, receptive field size or model depth, is more critical. To determine the answer, we focus on the single image super-resolution task, and propose a strategy based on dilated convolution to investigate how the two factors affect the performance. Our findings from exhaustive investigations suggest that single image super-resolution is more sensitive to the changes of receptive field size than to model depth variations, and that the model depth must be congruent with the receptive field size to produce improved performance. These findings inspire us to design a shallower architecture which can save computational and memory cost while preserving comparable effectiveness with respect to a much deeper architecture.

Finally, we study the general non-blind image deconvolution problem. It is observed in practice that by using existing deconvolution techniques, the residual between the sharp image and the estimation is highly dependent on both the sharp image and the noise. These techniques require the construction of different restoration models for different blur kernels and noises, inducing low computational efficiency or highly redundant model parameters. Thus, for general purposes, we propose a method by designing a very deep convolutional neural network which can handle different kernels and noises, while preserving high effectiveness and efficiency. Instead of directly outputting the deconvolved results, the model predicts the residual between a pre-deconvolved image and the corresponding sharp image, which can make the training easier and obtain restored images with suppressed artifacts.

Contents

List of Figures

List of Tables